REPOSITORIES SUPPORT PROJECT

0845 257 6860

Achieving Compliance with OpenAIRE Guidelines

support@rsp.ac.uk

Overview

As part of its Seventh Research Framework Programme (FP7) the European Commission is conducting a pilot initiative on Open Access to outputs from work that it funds. Publications in any of seven pilot areas of research must be made available as Open Access in either institutional or subject repositories. A central portal is being established to provide access to these publications but in order to allow their content to be indexed institutional repositories must comply with guidelines set out by the OpenAIRE project.

OpenAIRE and the OpenAIRE Guidelines

OpenAIRE[1,2] is a European Commission project to promote Open Access to peer reviewed outputs from FP7 funded projects. Special Clause 39 of the FP7 grant agreement states that publications in seven pilot research areas must be made Open Access through an institutional or subject repository within six months to a year from publication depending on the subject area.

OpenAIRE has set up a portal to use OAI-PMH to harvest metadata from repositories that contain FP7 publications. A set of guidelines[3] have been produced that repositories must meet in order for the metadata to be harvested correctly. These build on requirements from a previous project, DRIVER[4]. The remainder of this document will summarise the main points of the OpenAIRE guidelines and provide information on achieving compliance with the most popular institutional repository software in the UK.

OAI-PMH

The OpenAIRE portal will retrieve data from compliant repositories using the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)[5]. This is a machine-to-machine protocol for obtaining metadata regarding items contained in a repository. Records can be obtained from the repository as a whole or broken down into sets, which usually consist of the browsing hierarchy of the repository.

Requirements for Compliance – Key Points

The OpenAIRE guidelines document provides a full description of the requirements for repositories to comply with the FP7 Open Access pilot. The requirements can be broken down into those that are mandatory and those that are recommended.

Mandatory components

The mandatory parts of the guidelines can be summarised as follows:

- **Mandatory metadata elements**: The following elements must be contained in the Dublin Core (DC) metadata provided via the OAI-PMH interface of the repository: *title*, *creator*, *date* (publication), *type*, *identifier*, *relation* (Project ID – see below), *rights*
- **Inclusion of Project ID**: The DC relation element must be used to specify the ID of the FP7 project that funded the research. The ID should follow a specific format as defined in the guidelines
- **Metadata formats**: In addition to the project ID the following metadata elements should follow defined formats or draw from controlled vocabularies:
  - *type*. This element should consist of one of the types defined in the DRIVER guidelines 2.0, page 69

o   *rights.* This element should consist of values defined in the OpenAIRE Guidelines 1.1, page 8
- **EC_fundedresources set**: All *peer reviewed* outputs of FP7 projects in the 7 pilot subject areas should be available through an OAI-PMH set called EC_fundedresources with set specification ec_fundedresources

In addition the guidelines make various recommendations of which these are the highlights:

- **Embargo date**: Where the rights element indicates that an embargo exists a date field should be included that contains the end date of the embargo. This should be formatted as defined in the guidelines
- **Names**: all names, e.g. creator, should be in the format "last name, first name(s)"
- **Languages**. Where language is specified, 3 letter language codes in ISO 639-3 format should be used

The OpenAIRE project is currently working on a repository validator and registration system on the OpenAIRE web site. Once this is available managers should ensure that their site is registered with OpenAIRE so that their EC_fundedresources set may be harvested by the project.

Compliance for DSpace 1.6.2+

The OpenAIRE partners at FECYT, Spain have produced a quick start guide[6] for compliance with the OpenAIRE guidelines. While this contains a general summary of the requirements it is aimed mainly at DSpace repositories.

The Portuguese RCAAP projects has produced an add-on[7] that extends the OAI-PMH capability of DSpace that simplifies the process of creating the EC_fundedsesources set.

Compliance for EPrints 3.2+

The Repositories Support Project has developed an add-on for EPrints 3.2[8] that can aid repository managers in achieving compliance with the guidelines. It automates the process of creating the EC_Fundedresources set and adds the additional required metadata fields to the deposit workflow.

We would advise that EPrints administrators also read the quick start guide developed by FECYT[6] in addition to the guidelines themselves. While this document is DSpace centric it does also provide a useful overview of the steps required to meet the guidelines.

Limitations

Some limitations exist in the way in which EPrints creates custom sets and so the add-on will not guarantee compliance with the guidelines without additional work. EPrints custom set definitions only allow a single filter to be used to select items for the set. To create the EC_fundedresources set this filter find all records with a specific flag in the metadata. However, this would not exclude metadata only records or enforce the existence of a Rights field, for example. A global filter could be added to ensure that only full text records are selected but this would apply to *all* OAI-PMH sets, not just EC_fundedresources, meaning that the repository would only produce OAI-PMH data for records containing full text. As this may not be desirable for all repositories it was not included as part of the add-on. Therefore, this and other checks may need to be enforced using manual processes.

OpenAIRE 'orphan repository'

The OpenAIRE project provides an 'orphan repository'[9] that can be used to archive the outputs of FP7 projects. This is provided specifically for projects that do not have access to a compliant institutional or subject repository.

Further information:

[1] **OpenAIRE**, http://www.openaire.eu. UK/Western Europe contact: openaire@nottingham.ac.uk

[2] **OpenAIRE page at the Centre for Research Communications**
http://crc.nottingham.ac.uk/projects/openaire/

[3] **OpenAIRE Guidelines v1.1**
http://www.openaire.eu/index.php?option=com_content&view=article&id=207&lang=en

[4] **DRIVER Guidelines v2.0** http://www.driver-support.eu/managers.html

[5] **Open Archives Initiative Protocol for Metadata Harvesting**
http://www.openarchives.org/OAI/openarchivesprotocol.html

[6] **Repository Managers and Administrators Quick Guide to facilitate compatibility with OpenAIRE Guidelines**
http://www.openaire.eu/index.php?option=com_attachments&task=download&id=87%E2%8C%A9%3Den

[7] **RCAAP OAI Extended add-on for DSpace 1.6.2+**
http://projecto.rcaap.pt/index.php?option=com_remository&Itemid=2&func=fileinfo&id=321&lang=pt

[8] **RSP OpenAIRE Compliance add-on for EPrints 3.2+** http://files.eprints.org/649/, see
http://wiki.eprints.org/w/Files/OpenAIRE_Compliance_Plug-in for documentation

[9] **OpenAIRE orphan outputs repository** http://openaire.cern.ch/